FEBS Letters 583 (2009) 1422–1426

journal homepage: www.FEBSLetters.org

High throughput sequencing of microRNAs in chicken somites

Tina Rathjen [a,1], Helio Pais [b,1], Dylan Sweetman [a], Vincent Moulton [b], Andrea Munsterberg [a,*], Tamas Dalmay [a,*]

[a] School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK
[b] School of Computing Science, University of East Anglia, Norwich NR4 7TJ, UK

ARTICLE INFO

Article history:
Received 11 September 2008
Revised 25 February 2009
Accepted 19 March 2009
Available online 27 March 2009

Edited by Shou-Wei Ding

Keywords:
Deep sequencing
microRNA
Chicken
Somite

ABSTRACT

High throughput Solexa sequencing technology was applied to identify microRNAs in somites of developing chicken embryos. We obtained 651 273 reads, from which 340 415 were mapped to the chicken genome representing 1701 distinct sequences. Eighty-five of these were known microRNAs and 42 novel miRNA candidates were identified. Accumulation of 18 of 42 sequences was confirmed by Northern blot analysis. Ten of the 18 sequences are new variants of known miRNAs and eight short RNAs are novel miRNAs. Six of these eight have not been reported by other deep sequencing projects. One of the six new miRNAs is highly enriched in somite tissue suggesting that deep sequencing of other specific tissues has the potential to identify novel tissue specific miRNAs.

1. Introduction

MicroRNAs (miRNAs) are non-coding RNA molecules 20–23 nucleotide (nt) in length. They regulate the expression of genes by guiding the RNA Induced Silencing Complex (RISC) to a target sequence, which is usually located at the 3′ untranslated region (UTR) of mRNAs [1]. Interaction between RISC and target sites results in translational suppression of the mRNA and often leads to decreased mRNA stability [2–6]. There are 678 known miRNA loci in the human genome (miRBase release 11.0) [7] and each miRNA is expected to target hundreds of genes illustrating that miRNA-mediated suppression of protein synthesis represents an extensive layer of gene expression regulation [8].

Traditional sequencing of cDNA libraries made from short RNAs is very laborious and preferentially identifies abundant miRNAs [1]. Although some miRNAs are very highly expressed, others are present at a low level in a tissue because they are expressed only in a small number of cells [9]. Recent developments in high throughput sequencing technology allowed identification of hundreds of thousands/millions of short RNAs in a sample in various organisms [10–18]. These technologies can be used to find miRNAs expressed at very low level or in a small number of cells.

Chicken is an established model organism to study vertebrate development, largely due to the accessibility of the embryos in the egg and the ease with which they can be manipulated. The developmental stages of the embryo are well defined [19] and experimental techniques include classic and molecular approaches. Several genomic resources are now available, with EST data bases and the completed genome of *Gallus gallus* (red jungle fowl) [20]. The number of chicken miRNAs in miRBase 11.0 (149) is significantly lower than the number of human (678), mouse (472) or zebrafish (337) miRNAs suggesting that many chicken miRNAs are yet to be discovered. In fact, several groups reported recently novel chicken miRNAs by traditional [21,22] or deep sequencing [17,18] from whole embryos [21,18], primary fibroblast cells [17] or MDV-transformed lymphoblastoid T-cell line MSB-1 [22]. However, the relatively small overlap between the new miRNAs described by these reports indicates that these libraries are not saturated and specific tissues contain other unidentified miRNAs.

We were particularly interested in characterizing short RNAs in developing somites. Somites are transient mesodermal structures which form on either side of the neural tube in the embryonic midline. They are generated sequentially from the unsegmented presomitic mesoderm in an anterior to posterior (head to tail) direction [23]. The ventral part of somites gives rise to the sclerotome containing precursor cells for cartilage and bone and dorsal somite cells develop into the dermomyotome, which contains progenitor cells for skeletal muscle and skeletal muscle stem cells. Recent studies have confirmed that somites also give rise to connective

* Corresponding authors. Fax: +44 1603 592250.
E-mail addresses: a.munsterberg@uea.ac.uk (A. Munsterberg), t.dalmay@uea.ac.uk (T. Dalmay).
[1] These authors contributed equally to this work.

0014-5793/$36.00
doi:10.1016/j.febslet.2009.03.048

tissue, such as tendon cells, and endothelial and smooth muscle cells [24]. A few miRNAs were shown to be specifically expressed in somites, most notably the muscle specific microRNAs, miR-1, 206 and 133 [25–27], but it was never tested whether there are other somite specific miRNAs. Here we describe the deep sequencing of short RNAs isolated from dissected chicken somites and the experimental validation of eight novel miRNAs.

2. Materials and methods

2.1. Cloning of chicken miRNAs and Northern blot analysis

The overlying ectoderm and underlying endoderm were dissected using a tungsten needle. Dissected tissue white leghorn chicken embryos was disaggregated and RNA extracted using the miRVana kit (Ambion). Small RNA fraction between 19 and 24 nt was isolated from 15% denaturing polyacrylamide gel and 15 μg was ligated to Solexa adaptors (Illumina) without de-phosphorylating and re-phosphorylating. The short RNAs were converted to DNA by RT-PCR following the Illumina protocol and the DNA was sequenced on a Solexa machine (Illumina). Fifteen micrograms of RNA extracted from dissected somite tissue was analysed by Northern blot as described by Pall et al. [30]. Membranes were hybridised with non-modified 5′ labelled oligonucleotides complementary to the miRNA candidates at 37° using ULTRAhyb-Oligo buffer (Ambion) overnight.

2.2. Sequence analysis

Chicken genome was downloaded from Ensembl (http://www.ensembl.org/, assembly WASHUC2, database version 49.2). Sequence reads were mapped to the genome using PatMaN [31]. To generate the microRNA candidates MiRCat software (http://srna-tools.cmp.uea.ac.uk/) [28] was used. Sets of microRNAs for chicken were downloaded from miRBase (http://microrna.sanger.ac.uk/sequences/).

3. Results

3.1. Deep sequencing of short RNAs from chicken somites

Somites were dissected from 3-, 4- and 5-day old embryos and the extracted RNA was pooled. A cDNA library was generated from the short RNA (sRNA) fraction and sequenced using the Solexa technology (Illumina). We obtained 651273 reads representing 15976 distinct sequences (Table 1) (we use the term "distinct sequences" for sequences that are not identical to each other; the difference can be mismatches or shorter/longer sequences). The most abundant size class was 22 nt (43%) followed by the 21 nt reads (12%) (Fig. 1). 340415 reads (corresponding to 1701 sequences) mapped to the genome without any mismatch. Since almost half of the reads (310858) did not show a perfect match to the genome we tested whether these sequences were similar to the genome. 73011 (representing 4007 sequences) and 99425 (corresponding to 3469 sequences) reads could be mapped to the genome with one or two mismatches, respectively. Altogether, 78.7% of the reads (512851) mapped to the genome allowing up to two mismatches suggesting that the majority of the cDNA library derived from chicken sRNAs. The 1701 sequences showing perfect match to the genome were analysed further and 85 known chicken miRNAs were identified. The 85 sequences corresponded to 187975 reads and Additional file 1 shows the number of reads for each known miRNA.

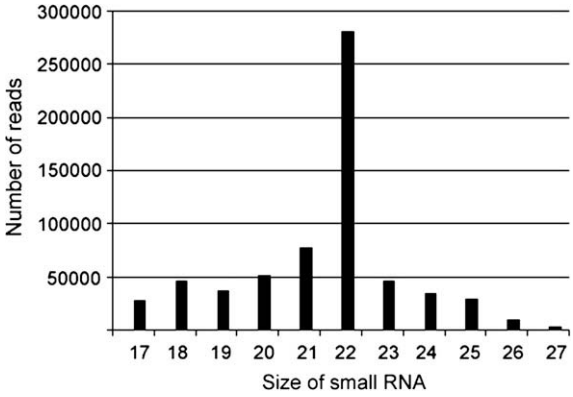

Fig. 1. Size distribution of sequenced short RNAs.

Table 1
Distribution of short RNAs. Number of reads and distinct sequences are shown of different categories of sequences.

	Number of reads	Number of distinct sequences
Total	651273	15976
Mapping to the genome without a mismatch	340415	1701
Mapping to the genome with one mismatch	73011	4007
Mapping to the genome with two mismatches	99425	3469
Known miRNAs	187975	85
New miRNA candidates	36656	42

3.2. Novel miRNAs

The remaining 1616 sequences that mapped to the genome but were not known miRNAs were analysed further using the miRCAT program we have developed earlier [28]. This program processes high throughput sequencing datasets, finds exact sequence matches in the genome, extracts flanking genomic sequences and tests whether they can be folded into hairpin structures where the cloned sequences are located in one of the stems of the hairpins. miRCAT identified 42 potential new miRNAs, which mapped to 45 loci in the genome (Table 2). Additional file 2 shows the predicted structures for all 45 loci. Twelve of these loci are located in introns of protein coding genes, three are in exons, 21 are in intergenic regions and nine are in un-annotated regions (Table 2). Eight of these loci have been predicted to encode for miRNAs according to ENSEMBL (sequences 10, 11, 17, 19, 27, 29, 30 and 31 in Table 2). We also compared these loci with the recently reported new potential chicken miRNA loci [17,18,21,22] (Table 2). Several sequences were cloned by more than one group but none of the sequences were found by all five groups but the limitation probably is the relatively small sample size of Shao et al. [21] and Yao et al. [22]. Twenty-three new potential miRNAs have not been reported before and one of these was mapped to four un-annotated loci.

3.3. Expression of novel miRNAs

Next we tested the accumulation of all 42 candidates in dissected somite RNA samples by Northern blot analysis (Fig. 2). Twenty-four oligonucleotide probes did not give any signal after 4-day exposures but the other 18 did. After further analysis of these 18 sequences by BLAST [29], 10 turned out to be variants of known miRNAs (Table 2). These sequences either contained one or two mismatches compared to known miRNAs or had

Table 2
Potential new miRNA genes. The first five columns show the in-house ID of sequences (No.), genomic location of the putative miRNAs (Chr: chromosome; Un: unknown chromosome; MT: mitochondrial genome), starting position on the chromosome (Start), the number of reads in our library (No. of reads) and the type of location of the miRNA gene (Location; 1: these miRNA genes are predicted in ENSEMBL; unknown location: these miRNA genes are situated in un-annotated regions of the genome). The sixth column (Found) shows the occurrences sequences in different libraries: S: somite (this study), F: fibroblast cells [17], E2: embryo [18], E3: embryo [21], T: T-cells [22]. The eighth column indicates the result of Northern blot analysis: negative result (−), a weak signal after a 4-day exposure (+) or a strong signal after a 2-h exposure (++). Some sequences confirmed by Northern blot are very similar to known miRNAs (either one or two mismatches or perfect match but has different length from the sequence in miRBase). These known miRNAs are indicated in the last column. Sequences in bold are confirmed by Northern blot and are not similar to any known miRNAs in miRBase. Please note that sequences 35, 38, 39 and 41 are identical and they could be derived from four different loci.

No.	Chr	Start	No. of reads	Location	Found	Sequence	Blot	miRNA
1	1	15746124	10	Intergenic region	S	**AGCGGCGCGGTAGGAGCA**	+	
2	1	1.7E + 08	5	Intergenic region	S	**CGGGAGGGGAGGGAGGGCGGG**	+	
3	1	1.04E + 08	25	Intergenic region	S	CGGGGATCGGGCGCGCCTCTCCGT	−	
4	1	62970293	20	Intergenic region	S	GTTTGGCTGTAGGCATGTGGGT	−	
5	1	1.31E + 08	25	Intron of predicted gene	S,E[2]	TGCATTGCGACGGGTTATATC	−	
6	1	1.04E+08	78	Intron of predicted gene	S	CGAGAAGACGGTCGAACTTGA	−	
7	2	1.07E+08	13	Intergenic region	S	TTCTGTAGACTGTTTGAC	−	
8	3	4348337	315	Exon of srf	S	ACGGGACAGTGCTGAAGACTAC	−	
9	3	23002020	14	Intergenic region	S,F,E[2]	TCCTGCAGAAGGTGCGGCT	−	
10	4	4049101	211	Intergenic region[1]	S,E[2],E[3],T	TAGCAGCACATCATGGTTTG	++	15b
11	4	92169345	1596	Intergenic region[1]	S,E[2]	TGAGAACTGAATTCCATGGACT	++	146a
12	5	33777672	32	Intron of CO041	S,E[2]	ACTAAGGACAGAGGAACGGAG	−	
13	5	45344529	26	Intergenic region	S	GTCGTCGGGATGGAGTTT	−	
14	5	60284030	26	Intergenic region	S,E[2]	CGGGGCGGCTGTGAGCTGAG	−	
15	5	26613484	27	Exon of INOC1	S	TCAGAAAAGGATATGAATTGTC	−	
16	6	34416337	39	Intron of ADAM12	S	TTAAGAGTAGGGATTCTGTTC	−	
17	8	29051925	233	Intergenic region[1]	S,E[2],T	GTACAGTACTGTGATAACTGAA	++	101
18	8	3838518	28	Intron of Q91015	S	TGGGTCCCGGCATGCTGCACT	−	
19	9	21966435	25	Intergenic region[1]	S,E[2],E[3]	GCGACCCATACTTGGTTT	+	551b
20	10	1811809	19	Intron of predicted gene	S	**TGCAGTGACGTCTCTTCCCC**	+	
21	10	11522320	16	Intergenic region	S	CAGGCGAGGGCGGGAGGGC	−	
22	14	4018538	111	Intergenic region	S	**CAGCAGGACTGGCTTTGTTACGA**	+	
23	14	857078	26	Intergenic region	S	GCGGAAGGACGGCGTCACTGG	−	
24	15	769601	25	Intergenic region	S,F,E[2]	ATCCCTTACTCACATGAG	−	
25	15	10171741	12	Intergenic region	S	CTGGAGGACACAGAGGCA	−	
26	15	1296957	65	Exon of DGCR8	S,F	TGGACGTTGGCTCTGGTGGTGA	−	
27	15	399864	47	Intergenic region[1]	S,E[2],E[3],T	TAGTGCAATATTGCTTATA	+	454
28	18	10554165	29	Intron of DMC1	S,E[2]	CGGCTTCTCGGTACCTGCGTT	−	
29	19	5352156	33	Intergenic region[1]	S,E[2],E[3],T	AGTTCTTCAGTGGCAAGCTT	+	22
30	19	7145042	199	Intron[1]	S,E[2],E[3],T	CAGTGCAATGTTAAAAGGGC	++	130a
31	20	8681835	181	Intergenic region[1]	S,E[2]	TAAGGCACGCGGTGAATG	++	124a
32	22	2685020	29	Intron of Q90715	S	AAGGTCCAACCTCACATGTCCT	−	
33	23	1213487	13	Intron of SCMH1	S	TGCGCTTTCTCATCCCGGC	−	
34	MT	9050	31	Unknown	S	AGCTAGAGAGAGGGGACAC	−	
35	Un	61344147	17	Unknown	S	GCAGGAGCGGGGCTCGGT	−	
36	Un	16238906	31	Unknown	S	ATGGGTCAAACGTTGACCAA	−	
37	Un	379349	28660	Unknown	S,E[2]	TACCCTGTAGATCCGAATTTGT	++	10b
38	Un	21596967	17	Unknown	S	GCAGGAGCGGGGCTCGGT	−	
39	Un	63782874	17	Unknown	S	GCAGGAGCGGGGCTCGGT	−	
40	Un	14731567	1596	Unknown	S,E[2]	TGAGAACTGAATTCCATGGACT	++	146a
41	Un	47015331	17	Unknown	S	GCAGGAGCGGGGCTCGGT	−	
42	Un	38326406	20	Unknown	S	**TCCCAGTGGAGCTCTGCAAGGACC**	+	
43	Z	44167547	1825	Intergenic region	S,F,E[2]	**AAAGGACGGAGGCGGCCCGCGC**	+	
44	Z	68816780	841	Intron of NCBP1	S	**ATGCAGAAGTGCACGGAAACAGCT**	++	
45	Z	34596495	31	Intron of XR_027016.1	S,F,E[2]	**TCCTTAACTCATGCCGCTGT**	+	

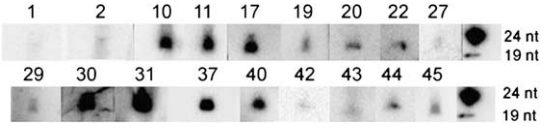

Fig. 2. Expression of new miRNAs. Total RNA from dissected somites were analysed by Northern blot. The numbers correspond to ID numbers in Table 2. Size markers (19 and 24 nt RNA oligonucleotides) are shown on the right.

slightly different length from miRNAs deposited in miRBase. The other 8 sequences were not similar to any known miRNAs and are bona fide new miRNAs. Interestingly, six of the eight new miRNAs were not found by other recently published high throughput sequencing projects. On the other hand, seven sequences that were present in our and at least in one other library were not confirmed by Northern blot analysis of somite RNA.

The expression of the six new miRNAs were further studied by dissecting the somites and hearts of 3-, 4- and 5-day old chicken embryos (the same stages we used for the sequencing). RNA was extracted from the somites and hearts and also from the remaining tissues of the dissected embryos. First we tested the accuracy of the dissection by hybridising miR-1 and miR-206 probes to the membranes. MiR-1 gave stronger signal in heart than in somite and miR-206 was expressed strongly in somite but not in the heart, as it was expected (Fig. 3) [25–27]. Both probes gave faint signals in tissues that contained the whole embryos except the heart and somites suggesting that either the dissection was not perfect or these muscle specific miRNAs have a low level of expression outside of the somites and heart. Nevertheless, this approach is appropriate to decide whether the new miRNAs are enriched in somites and/or heart. One of the six new miRNAs gave much stronger signal in somites than in heart of the rest of the embryo suggesting that this miRNA has a similar expression pattern to miR-206.

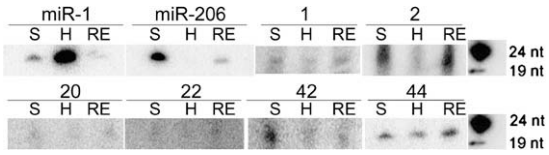

Fig. 3. Tissue specific expression of new miRNAs. Total RNA dissected from somites (S), hearts (H) and the rest of the embryos (RE) were analysed by Northern blot. The numbers correspond to ID numbers in Table 2. Probes to detect sequences 1, 2 and 42 were locked nucleic acid (LNA) containing primers. Size markers (19 and 24 nt RNA oligonucleotides) are shown on the right.

4. Discussion

Recent developments in high throughput sequencing technologies allow generation of large libraries of short RNAs [10–18]. The 454 technology gives the longest reads, currently 250–300 base pairs (bp), but significantly fewer reads than other techniques (about 400 000 per sample) [16,17]. The Solexa platform (Illumina) can identify up to 35 bp sequences and yields 1–3 million reads per sample [18]. Massively parallel sequencing (MPSS) gives only reads of 17 bp but even more reads than Solexa [12]. Since miRNAs are only 21–23 nt sequences, even MPSS can identify them reliably, although this technology does not give full length sequences therefore it is not the best option for miRNA discovery but rather for profiling known miRNAs. We chose the Solexa platform to test whether new miRNAs can be found in dissected somites of chicken embryos.

We have got 651 273 reads from chicken somites using Solexa but only 52% of them matched the genome perfectly. Another 26% of the reads could be mapped to the genome with one or two mismatches and many of these reads showed almost perfect match to known miRNAs from other species. This suggests that these reads were generally good sequences but due to some reason did not match the genome perfectly. There are at least two possible explanations for this. It is possible that there are single nucleotide polymorphisms (SNPs) between the sequenced genome (red jungle fowl) and the species we have used (white leghorn), however, at the moment there is not information about the extent of SNPs. It is also possible that Solexa generated sequences contain mistakes. Unfortunately we could not compare the ratio of reads perfectly matching the genome with the other available chicken short RNA library generated by Solexa [18] because this information was not given in the report.

After a bioinformatic analysis of the sequence reads we predicted 42 new miRNAs. We tested all 42 sequences by Northern blot analysis and accumulation of 18 was confirmed. We cannot rule out that the 24 negatives are true miRNAs because it is possible that they accumulate as 21–23 nt RNA but in a very small number of cells within the somites. In this case Northern blot analysis may not be sensitive enough to detect their expression. Recently, two groups reported new chicken miRNAs identified by deep sequencing of short RNAs from either whole embryos [18] or embryonic fibroblast cells [17]. Glazov et al. [18] used the Solexa platform and generated 9.5 million reads using pooled RNA from 5, 7 and 9-day old chicken embryos and identified 449 candidate miRNAs, although none of these were confirmed by other techniques. Burnside et al. [17] obtained 125 463 reads by 454 technology from fibroblast cells and found 63 new miRNA candidates. Five of these were confirmed by Northern blot analysis. One of these five (ID 58) is the same as one of our confirmed new miRNAs (ID 45, Table 2). Surprisingly, seven sequences, we and at least one other group found, were not confirmed by Northern blot. This suggests that some of the Northern negative sequences do accumulate as 21–23 nt RNA species but are below the detection level of Northern blot analysis either due to low expression level or because of a very restricted expression pattern. Another observation regarding the detection of miRNA expression by deep sequencing and Northern blot analysis is that no correlation was found between the number of reads and the strength of the hybridisation signal of different sequences. For example, probe 31 gave a stronger signal than probe 37 (Fig. 2) in spite of sequence 31 was sequenced 158 times less than sequence 37 (181 and 28 660 reads, respectively; Table 2). At the moment it is not clear why there is this difference. It is possible that the number of reads does not always reflect the accumulation level of individual sequences. It is also possible that some sequences give stronger signal due to higher GC content than other sequences even if they are accumulated at a lower level (although probes 31 and 37 had almost identical GC content). However, it is also worth considering that we [16] and others [11] found good correlation between the number of reads and strength of signal of specific sequences in different samples. For example if one sequence was cloned many times in one sample and few times in another sample, Northern blot analysis usually confirms the difference in expression level of that particular sequence between the two samples [11,16].

Ten of the 17 Northern positive probes detected sequences that were similar to known miRNAs, although miR-22 and miR-551 have not been found in chicken, only in other vertebrates. Eight sequences with confirmed expression are new miRNAs and six of these were not sequenced by others. Northern blot analysis of dissected tissues (somites, heart and the rest of the embryo) suggests that one of the six new miRNAs is specifically expressed in somites.

Acknowledgements

The work was supported by the European Commission (FP6 Integrated Projects SIROCCO LSHG-CT-2006-037900 to T.D.) and by the BBSRC (Ref. 016444 to A.M. and T.D.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2009.03.048.

References

[1] Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism and function. Cell 116, 281–297.
[2] Wightman, B., Ha, I. and Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. Cell 75, 855–862.
[3] Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R. and Pasquinelli, A.E. (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell 122, 553–563.
[4] Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E. and Filipowicz, W. (2005) Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science 309, 1573–1576.
[5] Petersen, C.P., Bordeleau, M.E., Pelletier, J. and Sharp, P.A. (2006) Short RNAs repress translation after initiation in mammalian cells. Mol. Cell 21, 533–542.
[6] Eulalio, A., Huntzinger, E. and Izaurralde, E. (2008) GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. Nat. Struct. Mol. Biol. 15, 346–353.
[7] Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R. and Bateman, A. (2005) Rfam annotating non-coding RNAs in complete genomes. Nucleic Acids Res. 33 (Database issue), D121–D124.
[8] Grimson, A., Farh, K.K., Johnston, W.K.K., Garrett-Engele, P., Lim, L.P. and Bartel, D.P. (2007) MiRNA targeting specificity in mammals: determinants beyond seed pairing. Mol. Cell 27, 91–105.
[9] Chang, S., Johnston Jr., R.J., Frokjaer-Jense, C., Lockery, S. and Hobert, O. (2004) MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. Nature 430, 785–789.
[10] Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C. and Green, P.J. (2005) Elucidation of the small RNA component of the transcriptome. Science 309, 1567–1569.
[11] Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J. and Jacobsen, S.E. (2006) Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat. Genet. 38, 721–725.

[12] Lu, C., Kulkarni, K., Souret, F.F., MuthuValliappan, R., Tej, S.S., Poethig, R.S., Henderson, I.R., Jacobsen, S.E., Wang, W., Green, P.J., et al. (2006) MicroRNAs and other small RNAs enriched in the Arabidopsis RNAdependent RNA polymerase-2 mutant. Genome Res. 16 (10), 276–1288.
[13] Berezikov, E., Thuemmler, F., van Laake, L., Kondova, I., Bontrop, R., Cuppen, E. and Plasterk, R. (2006) Diversity of microRNAs in human and chimpanzee brain. Nat. Genet. 38, 1375–1377.
[14] Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H. and Bartel, D.P. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell 127, 1193–1207.
[15] Burnside, J., Bernberg, E., Anderson, A., Lu, C., Meyers, B.C., Green, P.J., Jain, N., Isaacs, G. and Morgan, R.W. (2006) Marek's disease virus microRNAs map to meq and LAT. J. Virol. 80, 8778–8786.
[16] Moxon, S., Jing, R., Szittya, G., Schwach, F., Rusholme Pilcher, R.L., Moulton, V. and Dalmay, T. (2008) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. Genome Res. 18 (10), 1602–1609.
[17] Burnside, J., Ouyang, M., Anderson, A., Bernberg, E., Lu, C., Meyers, B.C., Green, P.J., Markis, M., Isaacs, G., Huang, E. and Morgan, R.W. (2008) Deep sequencing of chicken microRNAs. BMC Genomics 9, 185–194.
[18] Glazov, E.A., Cottee, P.A., Barris, W.C., Moore, R.J., Dalrymple, B.P. and Tizard, M.L. (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. Genome Res. 18 (6), 957–964.
[19] Hamburger, V. and Hamilton, H.L. (1951) A series of normal stages in the development of the chick embryo. J. Morphol. 88, 49–92.
[20] Hillier, L.W., Miller, W., Birney, E., Warren, W., Hardison, R.C., Ponting, C.P., Bork, P., Burt, D.W., Groenen, M.A., Delany, M.E., et al. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432, 695–716.
[21] Shao, P., Zhou, H., Xiao, Z.D., He, J.H., Huang, M.B., Chen, Y.Q. and Qu, L.H. (2008) Identification of novel chicken microRNAs and analysis of their genomic organization. Gene 418 (1–2), 34–40.
[22] Yao, Y., Zhao, Y., Xu, H., Smith, L.P., Lawrie, C.H., Watson, M. and Nair, V. (2008) MicroRNA profile of Marek's disease virus-transformed T-cell line MSB-1: predominance of virus-encoded microRNAs. J. Virol. 82, 4007–4015.
[23] Aulehla, A. and Herrmann, B.G. (2004) Segmentation in vertebrates: clock and gradient finally joined. Genes Dev. 18, 2060–2067.
[24] Smith, T.G., Sweetman, D., Patterson, M., Keyse, S.M. and Munsterberg, A. (2005) Feedback interactions between MKP3 and ERK MAP kinase control scleraxis expression and the specification of rib progenitors in the developing chick somite. Development 132, 1305–1314.
[25] Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S. and Plasterk, R.H. (2005) MicroRNA expression in zebrafish embryonic development. Science 309, 310–311.
[26] Darnell, D.K., Kaur, S., Stanislaw, S., Konieczka, J.K., Yatskievych, T.A. and Antin, P.B. (2006) MicroRNA expression during chick embryo development. Dev. Dyn. 235, 3156–3165.
[27] Sweetman, D., Rathjen, T., Jefferson, M., Wheeler, G., Smith, T.G., Wheeler, G.N., Munsterberg, A. and Dalmay, T. (2006) FGF-4 signaling is involved in mir-206 expression in developing somites of chicken embryos. Dev. Dyn. 235, 2185–2191.
[28] Moxon, S., Schwach, F., MacLean, D., Dalmay, T., Studholme, D.J. and Moulton, V. (2008) A toolkit for analysing large-scale plant small RNA datasets. Bioinformatics 24 (19), 2252–2253.
[29] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
[30] Pall, G.S., Codony-Servat, C., Byrne, J., Ritchie, L. and Hamilton, A. (2007) Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. Nucleic Acids Res. 35 (8), e60.
[31] Prufer, K., Stenzel, U., Dannemann, M., Green, R.E.E., Lachmann, M. and Kelso, J. (2008) PatMaN: rapid alignment of short sequences to large databases. Bioinformatics 24 (13), 1530–1531.